Texture to the Rescue: Practical Paper Fingerprinting based on Texture Patterns

Ehsan Toreini[1], Siamak F. Shahandashti[2], Feng Hao[1]

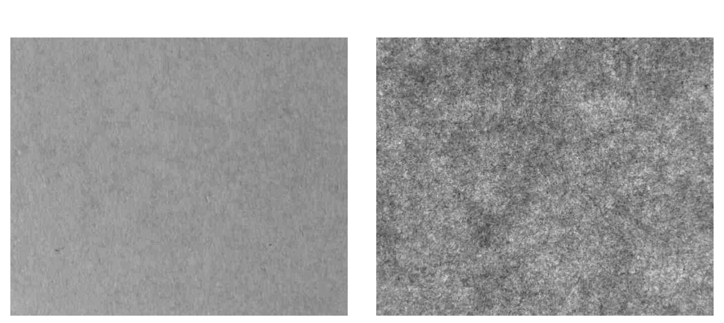
[1]Newcastle University, [2]University of York

Abstract

In this paper, we propose a novel paper fingerprinting technique based on analyzing the translucent patterns revealed when a light source shines through the paper. These patterns represent the inherent texture of paper, formed by the random interleaving of wooden particles during the manufacturing process. We show these patterns can be easily captured by a commodity camera and condensed into to a compact 2048-bit fingerprint code. Prominent works in this area (Nature 2005, IEEE S&P 2009, CCS 2011) have all focused on fingerprinting paper based on the paper "surface". We are motivated by the observation that capturing the surface alone misses important distinctive features such as the non-even thickness, the random distribution of impurities, and different materials in the paper with varying opacities. Through experiments, we demonstrate that the embedded paper texture provides a more reliable source for fingerprinting than features on the surface. Based on the collected datasets, we achieve 0% false rejection and 0% false acceptance rates. We further report that our extracted fingerprints contain 807 degrees-of-freedom (DoF), which is much higher than the 249 DoF with iris codes (that have the same size of 2048 bits). The high amount of DoF for texture-based fingerprints makes our method extremely scalable for recognition among very large databases; it also allows secure usage of the extracted fingerprint in privacy-preserving authentication schemes based on error correction techniques.

1. Motivation

- Unlike previous works that measure the paper surface characteristics, we propose to fingerprint a paper sheet based on measuring the paper texture patterns.
- We capture the texture by putting a light source on one side of the paper and using a commodity camera to take a photograph on the other side.
- We design an efficient paper fingerprinting algorithm, and carry out experiments to show that our method is reliable, accurate and inexpensive to deploy in practice.
- We conduct further experiments to demonstrate that our method is robust against: (a) non-ideal photo capturing settings when the paper is rotated and the light source is changed, and (b) non-ideal paper handling situations such as crumpling, soaking, heating and pen scribbling.
- The figure below shows the visual difference between the paper surface and the paper texture.



(a) Paper Surface (b) Paper Texture

Figure: The surface and texture of the same area of a paper sheet as captured by a camera based on a) reflective and b) transmissive light.

2. Paper Texture

1. Two different light sources are used in our prototype and experiments.

(a) Overhead projector (b) Light box

Figure: Variety of Light Sources in Our Experiments.

2. All images are taken by using a standard off-the-shelf camera (Panasonic DMC-FZ72).
3. A macro ring flash is mounted on the camera to keep a constant distance to the paper surface.

4. A box is printed on the paper surface and only the texture within this box is captured and analysed.

Figure: Using a commodity camera to capture the paper texture

3. Texture Analysis

1. A digital photograph of the paper sheet is taken with a backlit light source.
2. Then, a series of preparation operations are performed to align and resize the original image.

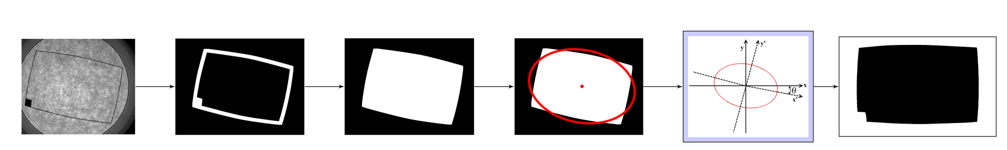


Figure: Step-by-step rotation recognition process in the preparation phase. The last step produces a mask that distinguishes the pixels containing reliable information suitable for feature extraction from the pixels containing unreliable information.

3. In the texture analysis phase, a 2-D Gabor filter is utilized to extract textural information from the captured image.
4. Our paper fingerprint extraction method generates a 2048-bit binary string, called the *paper fingerprint*.
5. Once paper fingerprints are in the binary string format, they can be compared based on the fractional Hamming distance.

4. Evaluation Results

1. Our hypothesis is that textural patterns revealed by the transmissive light contain richer features than the paper surface shown by the reflective light.
2. To validate the hypothesis, we selected 10 common A4 (210×297 mm) paper sheets with grammage 80 g/m^2, and took 10 photos for each sheet to compare the surface and texture patterns.

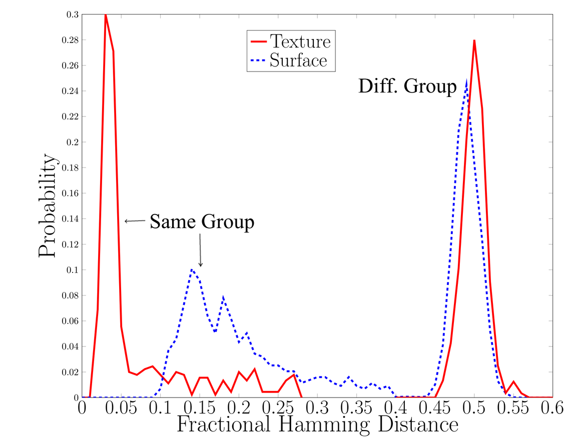


Figure: Hamming distance distributions for surface and texture.

3. Our main dataset comprises 1000 samples collected by taking 10 photos of each of 100 different paper sheets.
4. We use typical office paper sheets of size A4 (210mm × 297mm) with grammage of 80 g/m^2. All the sheets were from the same pack with the same brand.

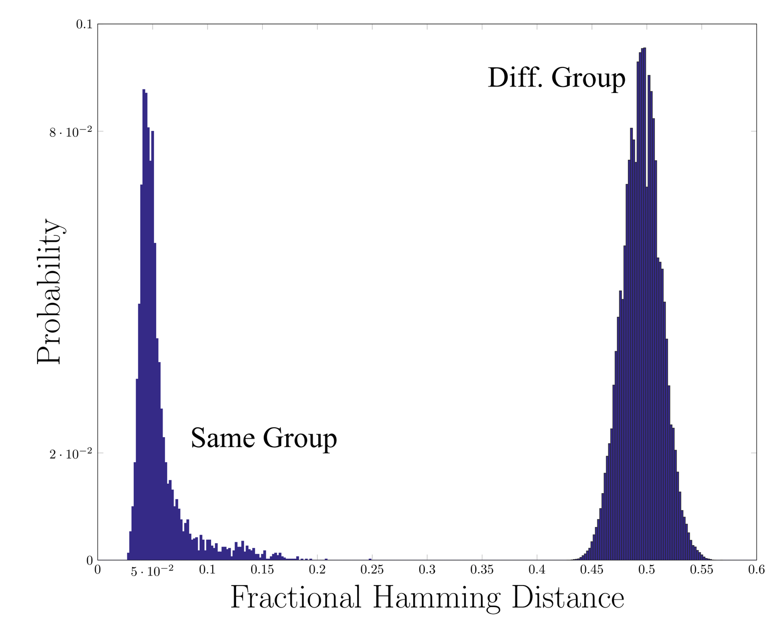
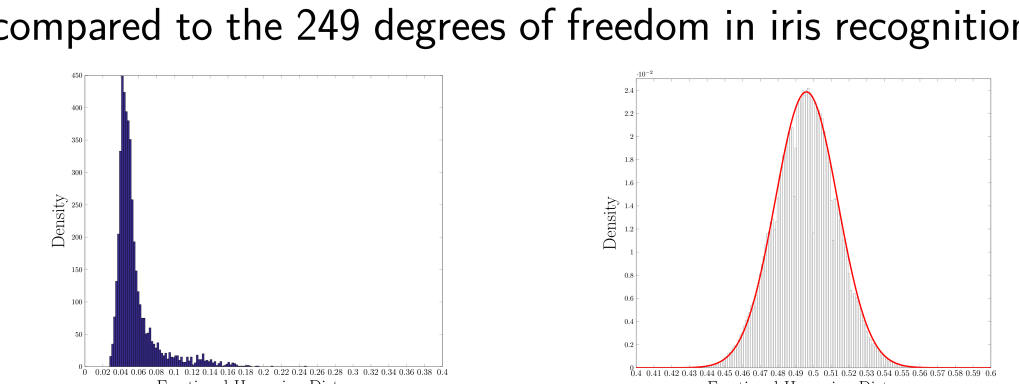


Figure: Hamming distance distributions in the benchmark dataset.

5. Decidability is $d' \approx 21$, which compares favourably to $d' \approx 14$ for iris recognition.
6. The number of degrees of freedom is $N = 807$, which means the entropy of the extracted fingerprints is 807 bits out of a total of 2048 bits. This is much better as compared to the 249 degrees of freedom in iris recognition.



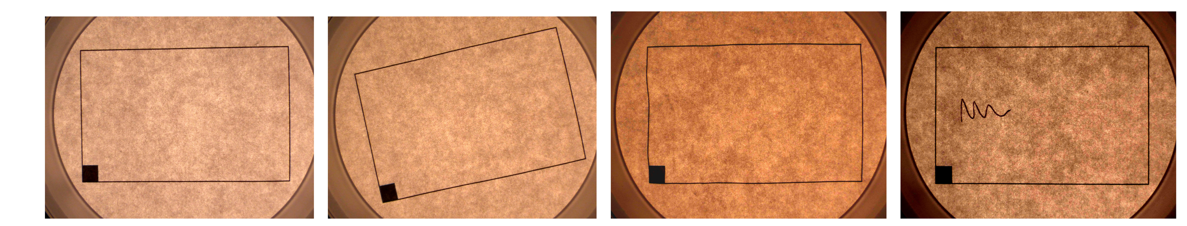
(a) Histogram of same-group HDs with $\mu = 0.056$, $\sigma = 0.024$

(b) Histogram of different-group HDs with a binomial curve with $N = 807$, $\mu = 0.495$, $\sigma = 0.018$

Figure: Histograms of Hamming distances in the benchmark dataset.

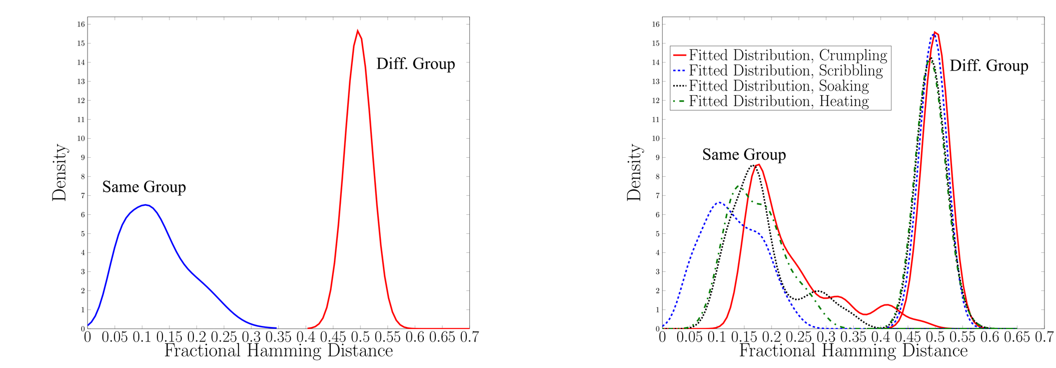
5. Robustness Evaluations

1. Impact of Non-Ideal Data Collection: Photo Rotation
2. Impact of Non-Ideal Paper Handling: Crumpling, Scribbling, Soaking, Heating
3. Impact of a Different Light Source

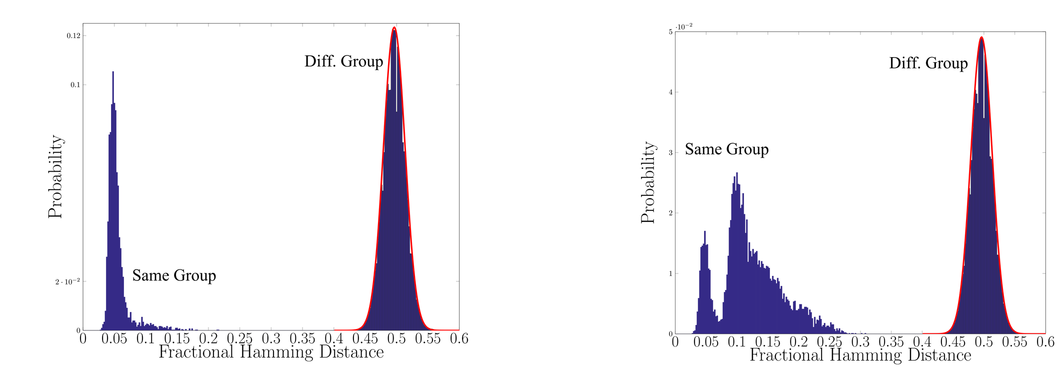


(a) Benchmark (b) Rotated (c) Crumpled (d) Scribbled

Rate	Ideal Value	Benchmark Dataset	Rotated	Crumpled	Scribbled	Soaked	Heated	Mixed Light
FAR	0%	0%	0%	0%	0%	0%	0%	0%
FRR	0%	0%	0.32%	3.2%	0%	0%	0%	0%



(e) Fitted distributions under rotation. (f) Fitted distributions under non-ideal paper handling.



(g) Hamming distance distributions for the light box dataset, plus the binomial curve with $N = 846$, $\mu = 0.496$, $\sigma = 0.017$

(h) Hamming distance distributions for the mixed light box and projector dataset, plus the binomial curve with $N = 836$, $\mu = 0.496$, $\sigma = 0.017$

6. Publication and media coverage

The paper is published in *ACM Transactions on Privacy and Security* 2017, and covered by the Economist, Wall Street Journal and many other media.

Contact Information

- https://toreini.github.io/projects/fingerprinting.html
- Email: e.toreini@ncl.ac.uk